Logique doxastique graduelle

Laverny Noël
laverny@irit.fr; http://clsl.chez-alice.fr

Institut de Recherche en Informatique de Toulouse

Résumé :
La modélisation des croyances est un sujet très important de l'intelligence artificielle. Nous présentons ici une logique modale permettant de raisonner sur des croyances plus ou moins fortes d'un agent sur le système. Nous définissons un langage permettant de gradualiser les croyances : de la croyance faible jusqu'à la conviction en passant par divers degrés. Nous donnons une axiomatique et une sémantique (complète et adéquate) basée sur les modèles de Kripke. Nous montrons ensuite que toute formule peut se réduire à une formule sans modalités imbriquées. Nous définissons alors des modèles numériques basés sur les fonctions conditionnelles ordinales de Spohn.

Mots-clés : Logique modale, croyances, croyances graduelles, fonction conditionnelle ordinale

Abstract:
Reasoning about beliefs is an important issue in artificial intelligence. We present here a modal logic allowing for reasoning about more or less strong beliefs held by an agent. We define a language for graded beliefs. We give then an axiomatics and a semantics based on Kripke models, together with a soundness and completeness result. We show that any formula can be reduced to a formula without nested modalities. We discuss an alternative semantics based on Spohn's ordinal conditional functions.

Keywords: Modal logic, beliefs, graded beliefs, ordinal conditional function

1 Introduction

La représentation de la connaissance (ou de la croyance) est un domaine déjà bien étudié de l'intelligence artificielle. Pour autant, les convictions que peut avoir un homme sur tel ou tel fait sont plus complexes qu'une simple connaissance (voire croyance). Celles-ci sont en constante évolution et remises en question. Nous faisons ici un premier pas dans la modélisation d'un "état épistémique" (l'ensemble des convictions d'un individu à un moment donné) en considérant des degrés de croyances rassemblés dans la logique doxastique graduelle.

La première personne à écrire sur la logique épistémique est le philosophe G.H. von Wright dans son livre "An essay in modal logic" en 1953 [18]. Son étude est uniquement axiomatique, sans utilisation de la sémantique des mondes possibles. Les travaux sur la logique épistémique de la plupart des philosophes suivants ont consisté à défendre ou attaquer les axiomes établis par von Wright. Ce n'est qu'après "l'invention" de la sémantique des logiques modales (modèles de Kripke), que, dans les années soixante, les sujets sur la logique épistémique fleurissent. Les champs d'applications sont multiples : la théorie des jeux, les systèmes distribués, la théorie de la décision, la négociation etc. On peut citer deux livres importants : "Reasoning about knowledge" [2] et "Epistemic Logic for AI and Computer Science" [13].

Dans la section 2 nous rappelons la définition du système modal de base permettant de raisonner sur connaissances et croyances. Puis nous définissons dans la section 3 un système graduel de croyances, nous en donnons les propriétés ainsi qu'une sémantique bâtie sur les OCFs. Nous discutons ensuite, dans la section 6, des travaux existant sur la "gradualisation" des croyances.

2 Connaissances et croyances

Les opérateurs habituellement utilisés sont "K" (comme Knowledge) et "B" (comme Belief). Leur interprétation est épistémique ($\mathbf{K}\Psi$ signifie "je sais que Ψ") et doxastique ($\mathbf{B}\Psi$ signifie "je crois que Ψ"). Les langages utilisés ici seront donc $\mathcal{L}(\mathtt{PS}, \{\mathbf{K}\})$ ($\mathtt{PS}$ désignant un ensemble fini de symboles propositionnels, $\mathcal{L}(\mathtt{PS}, \{\mathbf{K}\})$ est l'ensemble des formules bien formées à partir de $\mathtt{PS}$ et $\mathbf{K}$), $\mathcal{L}(\mathtt{PS}, \{\mathbf{B}\})$ ou $\mathcal{L}(\mathtt{PS}, \{\mathbf{K}, \mathbf{B}\})$ (l'agent pourra avoir des connaissances et des croyances). Ceux-ci sont assez expressifs et puissants pour raisonner sur les connaissances ou les croyances. Par exemple, il n'y a pas de restriction sur la *portée* des opérateurs, pouvant nous conduire à une formule du type $\mathbf{KB}p$ ("je sais que je crois p") ou $\mathbf{B}\neg\mathbf{K}p$ ("je crois que je ne sais pas p").

2.1 Le système KB

Définition 1 *Le système* KB *est la logique réunissant les axiomes du système* S5 *pour* $\mathbf{K}$*, les axiomes de* KD45 *pour* $\mathbf{B}$ *et les axiomes d'interactions suivants : A1.* $\mathbf{K}\Psi \rightarrow \mathbf{B}\Psi$ *et A2.* $\mathbf{B}\Psi \rightarrow \mathbf{KB}\Psi$

La connaissance doit donc être vraie et doit vérifier l'introspection positive et l'introspection négative. La croyance, elle, n'est pas forcement vraie, mais doit être consistante, et doit vérifier aussi l'introspection positive et l'introspection négative. La connaissance implique la croyance (A1) et l'agent est conscient de ses croyances (A2).
Tous ces axiomes constituent les bases de *notre* raisonnement sur connaissances et/ou croyances, ils sont assez simples et intuitifs et suffiront pour notre propos. Notons toutefois qu'il y a discussion possible sur l'ensemble des propriétés que *doivent* vérifier les connaissances, les croyances et les interactions entre les deux. Si nous faisons l'hypothèse appelée par Lenzen [11] *entailment property* ($\mathbf{K}\varphi \rightarrow \mathbf{B}\varphi$), ainsi que l'hypothèse de *conviction des croyances* ($\mathbf{B}\varphi \rightarrow \mathbf{BK}\varphi$), Lenzen, Lammarre et Shoham montrent ([11, 8]) que l'agent ne peut pas avoir de fausses croyances. Plusieurs solutions sont proposées :

- Dans [19], Voorbraak supprime l' "entailment property" et dans [5], Halpern l'affaiblit simplement aux formules *objectives* (formules où n'intervient aucune modalité $\mathbf{B}$ ni $\mathbf{K}$).
- Dans [4, 7, 14], les auteurs suppriment la propriété de "conviction des croyances". Il parlent alors de croyance faible, par opposition à la croyance forte, ou certitude.
- Dans [11, 8, 19], les auteurs affaiblissent les propriétés de la connaissance. La modalité $\mathbf{K}$ ne satisfait plus S5 mais une logique plus faible située entre S4 et S5 .

Le système que nous avons choisi correspond à des "croyances faibles" et possède la propriété suivante :

Proposition 1 *Toute formule de* $\mathcal{L}(\mathtt{PS}, \{\mathbf{K}, \mathbf{B}\})$ *est équivalente à une formule de profondeur modale inférieure ou égale à un*[1].

2.2 Quelques exemples

1. $\mathbf{K}(\mathrm{SurTable(livre)} \wedge \mathrm{Dehors(Lunette)})$ exprime : "l'agent sait que le livre est sur la table et les lunettes dehors"
2. Michel est parti à Toulouse. Je ne sais pas où Michel est parti, mais je sais qu'il est parti à Bordeaux ou à Toulouse, je crois plutôt Toulouse : $\mathbf{K}(\mathrm{Bordeaux} \vee \mathrm{Toulouse}) \wedge \mathbf{B}\mathrm{Toulouse}$. Je décide de lui téléphoner pour savoir ; quand je lui aurai téléphoné, j'aurai : $\mathbf{K}\mathrm{Bordeaux} \vee \mathbf{K}\mathrm{Toulouse}$. Je lui téléphone, et maintenant : $\mathbf{K}\mathrm{Toulouse}$.

3 La logique $\mathsf{KD45_G}$

L'étude précédente nous montre que la classe de modèles $\mathcal{KB}$ rallie expressivité et simplicité puisque toute formule

[1] une propriété analogue est démontrée dans [5] où l'axiome $\mathbf{K}\varphi \rightarrow \mathbf{B}\varphi$ ne s'applique qu'aux formules objectives et avec l'axiome de certitude des croyances $\mathbf{B}\varphi \rightarrow \mathbf{BK}\varphi$

a son équivalente sans modalités imbriquées. Ceci assure une taille raisonnable des formules et modère ainsi la complexité des problèmes de validité et de satisfiabilité.
Cependant, on ne peut pas, avec seulement les modalités $\mathbf{K}$ et $\mathbf{B}$, exprimer des degrés de conviction plus ou moins forts sur tel ou tel fait. Reprenons l'exemple précédent où je sais que Michel est parti à Bordeaux ($\mathtt{B}$) ou à Toulouse ($\mathtt{T}$), mais je crois plutôt Toulouse. Ceci s'exprime par : $\mathbf{K}(\mathtt{B} \vee \mathtt{T}) \wedge \mathbf{B}\mathtt{T}$. Imaginons que quelqu'un me dise : « Moi aussi, je crois qu'il est parti à Toulouse ». A ce moment là, ma croyance : "Michel est parti à Toulouse" s'est renforcée, ce que je ne peux exprimer avec $\mathcal{L}(\mathtt{PS}, \{\mathbf{K}, \mathbf{B}\})$. Ou encore, j'entends : « Ah, tu crois ? c'est plutôt Bordeaux, non ? ». Là, au contraire, ma croyance est affaiblie : si elle était très forte, j'y crois encore, si elle était moyenne, je n'ai maintenant plus aucune idée et si elle était faible, je crois maintenant que Michel est parti à Bordeaux. Se posent là deux problèmes :

1. Comment exprimer des croyances plus ou moins fortes ?
2. Comment réaliser la révision de celles-ci en les combinant avec d'autres croyances graduelles ?

Le premier point nécessite la création d'un nouveau langage et d'une nouvelle axiomatique. On pourra s'inspirer avantageusement de $\mathcal{L}(\mathtt{PS}, \{\mathbf{K}, \mathbf{B}\})$ pour cela[2].

3.1 Le langage des croyances graduelles

L'idée est de remplacer la modalité $\mathbf{B}$ par une famille de modalités $\mathtt{B} = \{\mathbf{B^i} : i \in \overline{\mathbb{N}}^*\}$[3]. $\mathbf{B^i}$ signifiant : « Je crois avec un *degré de conviction i* que... ». Plus le degré est grand, plus la croyance sera forte. A priori il n'y a pas lieu de se limiter à un degré de conviction maximum, on ira même jusqu'au degré infini ($\mathbf{B}^\infty$). Formellement le langage que nous allons utiliser est $\mathcal{L}(\mathtt{PS}, \{\mathbf{B^i} : i \in \overline{\mathbb{N}}^*\}) = \mathcal{L}^g_{\mathbf{B}}$.
$\mathbf{B^i}\varphi$ signifie : « l'agent croit "φ" avec un degré de conviction égal à i ». $\mathbf{B}^\infty\varphi$ signifie : « l'agent sait[4] "φ" ». Le langage permet bien sûr l'imbrication de modalités comme $\mathbf{B^1}\neg\mathbf{B^2}\varphi$ ou $\mathbf{B}^\infty\mathbf{B^3}(\varphi \vee \mathbf{B^1}\varphi)$

3.2 Le système $\mathsf{KD45_G}$

Définition 2 *Le système* $\mathsf{KD45_G}$ *est normal pour* $\{\mathbf{B^i} : i \in \overline{\mathbb{N}}^*\}$ *et possède les axiomes du système* $\mathsf{S5}$ *pour* $\mathbf{B}^\infty$*, les axiomes de* $\mathsf{KD45}$ *pour chaque* $\mathbf{B^i}$*,* $i \geqslant 1$ *et les schémas d'axiomes d'interaction suivants :*
A3. $\mathbf{B^i}\Psi \rightarrow \mathbf{B^j}\Psi$*, si* $i > j$
A4. $\mathbf{B^j}\Psi \rightarrow \mathbf{B}^\infty\mathbf{B^j}\Psi$

Autrement dit, $\mathbf{B}^\infty$ exprime la connaissance, c'est une modalité $\mathsf{S5}$. Chaque modalité $\mathbf{B^i}$, exprimant la croyance au degré i, est une modalité de $\mathsf{KD45}$. La "hiérarchie" prévue dans cette famille de modalités se retrouve dans l'axiome (A3). L'axiome (A4) est en quelque sorte l'homologue de l'axiome (A2), il exprime que, pour tout j, l'agent est conscient de ses croyances au degré j.

3.3 La classe de modèles $\mathcal{KD}45_\mathcal{G}$

Nous définissons maintenant une sémantique pour interpréter les formules de $\mathcal{L}^g_{\mathbf{B}}$.

Définition 3 $\mathcal{KD}45_\mathcal{G}$ *est la classe des modèles de Kripke* $M = (W, \mathcal{R}, \pi)$ *où*
– W *est l'ensemble des mondes possibles*
– $\mathcal{R} = \{R_i : i \in \overline{\mathbb{N}}^*\}$ *est un ensemble de relations d'accessibilité*

[2] Nous avons traité le deuxième point dans [10]

[3] $\overline{\mathbb{N}}^*$ désignant l'ensemble des entiers naturels privé de zéro et augmenté de l'infini (symbole ∞)

[4] Nous pourrions nous en tenir à : l'agent a la conviction que "φ", mais il est plus simple et plus raisonnable compte tenu des exemples que nous prendrons liés à la robotique, d'admettre que "tout ce dont est convaicu l'agent est vrai".

– $\pi : W \to \mathtt{PS} \to \{0,1\}$ *établit la valeur de vérité de chaque symbole propositionnel dans chaque monde.*

Chaque R_i pour $i < \infty$ (correspondant à la modalité $\mathbf{B^i}$) est sérielle, euclidienne et transitive et R_∞ (correspondant à la modalité $\mathbf{B^\infty}$) est une relation d'équivalence. De plus, pour tous $i, j : i > j$ on a : $R_\infty \circ R_j \subseteq R_j$ et $R_j \subseteq R_i$.

La valeur de vérité d'une formule Φ dans un monde w d'un modèle $M \in \mathcal{KD}45_\mathcal{G}$, notée $(M,w) \models \Phi$ est définie de manière classique pour $\Phi \in \mathtt{PS}$, $\neg$ et $\wedge$ et par : $(M,w) \models \mathbf{B^i}\Phi \Leftrightarrow (\forall v \in W\,(R_iwv \Rightarrow (M,w') \models \Phi))$.

La validité dans un modèle ($M \models \Phi$), la conséquence ($\Phi \models_{\mathcal{KD}45_\mathcal{G}} \Psi$) et l'équivalence ($\Phi \equiv_{\mathcal{KD}45_\mathcal{G}} \Psi$), sont définies de manière classique.

Proposition 2 $\mathsf{KD45_G}$ *est adéquat et complet vis-à-vis de $\mathcal{KD}45_\mathcal{G}$.*

Proposition 3 *Toute formule de $\mathcal{L}^g_\mathbf{B}$ est équivalente à une formule de profondeur modale inférieure ou égale à un.*

Exemple 1 *Pour toute formule objective φ, les formules suivantes sont des théorèmes :.*
– $\mathbf{B^1B^2}\varphi \leftrightarrow \mathbf{B^2}\varphi$
– $\mathbf{B^2}\neg\mathbf{B^1}\varphi \leftrightarrow \neg\mathbf{B^1}\varphi$
– $\mathbf{B^2}(\psi \vee \mathbf{B^3}\varphi) \leftrightarrow \mathbf{B^2}\psi \vee \mathbf{B^3}\varphi$
– $\mathbf{B^1}(\mathbf{B^1}\psi \wedge \mathbf{B^3}\varphi) \leftrightarrow \mathbf{B^1}\psi \wedge \mathbf{B^3}\varphi$

Remarque 1 *Intuitivement, on pourrait s'attendre à une simplification du type*

$$\models_{\mathcal{KD}45_\mathcal{G}} \mathbf{B^iB^j}\varphi \leftrightarrow \mathbf{B}^{\min(\mathbf{i},\mathbf{j})}\varphi \qquad (1)$$

On a bien $\models_{\mathcal{KD}45_\mathcal{G}} \mathbf{B^iB^j}\varphi \to \mathbf{B}^{\min(\mathbf{i},\mathbf{j})}\varphi$, mais la "réciproque", $\models_{\mathcal{KD}45_\mathcal{G}} \mathbf{B}^{\min(\mathbf{i},\mathbf{j})}\varphi \to \mathbf{B^iB^j}\varphi$, elle, n'est pas vraie. On peut même montrer que l'ajout dans l'axiomatique de cette propriété conduit à une dégénérescence des $\mathbf{B^i}$.

4 Les "OCFs" : des modèles pour $\mathsf{KD45_G}$

On va montrer dans ce paragraphe que toute formule consistante est satisfiable dans un modèle défini par une fonction ordinale[5] sur S (où S est l'ensemble de toutes les valuations possibles des symboles popositionnels : $2^{\mathtt{PS}}$). On introduit ces modèles parce qu'ils quantifient les degrés de conviction des croyances. Leur utilisation facilitera, par de simples opérations arithmétiques, le calcul des combinaisons, au niveau sémantique, de différentes croyances ([10]). Commençons par définir formellement les fonctions ordinales.

4.1 Les fonctions ordinales

Définition 4 *Une* fonction ordinale *sur l'ensemble des mondes (Ordinal Conditional Function [15] : OCF) est une fonction $\kappa : S \longmapsto \overline{\mathbb{N}}$, telle que $\min_{s \in S} \kappa(s) = 0$. κ peut s'étendre aux formules objectives par $\kappa(\varphi) = \min\{\kappa(s) \mid s \models \varphi\}$.*

Intuitivement, $\kappa(s)$ est le *degré d'exceptionnalité*[6] de s. En particulier
– $\kappa(s) = 0$ veut dire que s est un *monde normal* (un monde normal n'est exceptionnel à aucun degré)
– $\kappa(s) = 1$ veut dire que s est "simplement exceptionnel" ;
– $\kappa(s) = 2$ veut dire que s est "doublement exceptionnel" ;
– $\kappa(s) = \infty$ veut dire que s est impossible. Tout monde s tel que $\kappa(s) < \infty$ est appelé *monde possible* .

La *contrainte de normalisation* $\min_{s \in S} \kappa(s) = 0$ impose qu'il existe au moins un monde normal. L'OCF κ_{vide} est défini par $\kappa_{vide}(s) = 0$ pour tout s.

[5] inspirées de celles de Spohn [15]
[6] $\kappa(s)$ est usuellement interprété en termes de probabilités infinitésimales $\kappa(s) = k < +\infty$ signifie $prob(s) = o(\varepsilon^k)$, où ε est infiniment petit.

On utilisera la notation classique

$$\kappa = \begin{bmatrix} s_1 : n_1 \\ s_2 : n_2 \\ s_3 : n_3 \end{bmatrix}$$

qui signifie que $\kappa(s_i) = n_i$ pour $i = 1, 2$ et 3 et par convention que les autres états s de S, absents dans la représentation, sont impossibles i.e. $\kappa(s) = \infty$.

4.2 Vérité d'une formule doxastiquement interprétable dans une OCF

Toute combinaison booléenne de formules du type $\mathbf{B^i}\varphi$ est dite *doxastiquement interprétable*.

Définition 5 *Si φ est une formule objective et κ une OCF,*

$$\kappa \models \mathbf{B^i}\varphi \;\Leftrightarrow\; \kappa(\neg\varphi) \geqslant i$$

Ceci veut dire que $\kappa \models \mathbf{B^i}\varphi$ est vraie dès que tous les modèles s de $\neg\varphi$ sont exceptionnels au moins au degré i (on pourra dire aussi *i-exceptionnels*) i.e. sont tels que $\kappa(s) \geqslant i$. Ou encore, tous les états s' tels que $\kappa(s') < i$ (i.e. au plus $(i-1)$-exceptionnels) satisfont φ.
En particulier, $\mathbf{B^1}\varphi$ est vraie lorsque tout les états normaux (i.e. $\kappa(s) = 0$) satisfont φ, et $\mathbf{B^\infty}\varphi$ est vraie lorsque tous les états possibles (i.e. $\kappa(s) < \infty$) satisfont φ. Cette propriété est importante car on sera amené à utiliser ces formules pour représenter les croyances graduelles d'un agent. Une formule *doxastique positive* décrira ce que croit l'agent sur l'état du monde à tous les degrés de convictions possibles.

Exemple 2 *Soit* $\mathtt{PS} = \{a, b\}$ *et κ définie par :nous donn*[7]

$$\kappa = \begin{bmatrix} ab & : & 1 \\ a\bar{b} & : & 0 \\ \bar{a}b & : & 1 \\ \bar{a}\bar{b} & : & \infty \end{bmatrix}$$

Alors

- *$\kappa \models \mathbf{B^1}a \wedge \neg\mathbf{B^2}a$: l' agent croit a au degré 1 (car le (seul) monde normal, i.e. $a\bar{b}$, satisfait a), mais cette croyance n'est pas plus forte : a n'est pas cru au degré 2, car il y a $\neg a$-monde s tel que $\kappa(s) = 1$, il s'agit de $\bar{a}b$.*
- *$\kappa \models \mathbf{B^\infty}(a \vee b)$, car tous les états possibles (i.e., $a\bar{b}$, ab et $\bar{a}b$) satisfont $a \vee b$;*
- *$\kappa \models \neg\mathbf{B^1}b$, car le monde normal $a\bar{b}$ ne satisfait pas b.*

Proposition 4 *Soit Φ une formule doxastiquement interprétable, les deux propositions suivantes sont équivalentes : (1) Φ est satisfiable dans un $\mathcal{KD}45_\mathcal{G}$-modèle.
(2) Φ est satisfiable dans un OCF-modèle.*

4.3 Correspondance entre les formules et les OCFs

Définition 6 *Une formule est dite* normale doxastique positive *(on raccourcira avec le sigle NDP : Normal Doxastic Positive) lorsque $\Phi = \mathbf{B^\infty}\varphi_\infty \wedge \mathbf{B^n}\varphi_n \wedge \ldots \wedge \mathbf{B^1}\varphi_1$ où $\varphi_\infty, \varphi_1, \ldots, \varphi_n$ sont des formules objectives telles que pour tous i et $j > i$ on a, $\models \varphi_i \rightarrow \varphi_j$*

Quand on écrit une formule normale doxastique positive $\mathbf{B^\infty}\varphi_\infty \wedge \mathbf{B^n}\varphi_n \wedge \ldots \wedge \mathbf{B^1}\varphi_1$, on peut supprimer les sous formules $\mathbf{B^i}\varphi_i$ telle que $\varphi_{i+1} \equiv \varphi_i$, ainsi que les tautologies de la forme $\mathbf{B^i}\top$. Par exemple, $\mathbf{B^\infty}\top \wedge \ldots \wedge \mathbf{B^4}\top \wedge \mathbf{B^3}a \wedge \mathbf{B^2}a \wedge \mathbf{B^1}(a \wedge b)$ est simplement remplacée par $\mathbf{B^3}a \wedge \mathbf{B^1}(a \wedge b)$.
Les formules NDPs expriment tout ce que l'agent croit, elles sont satifaites par une famille d'OCFs (κ_j) possédant un plus petit élément que l'on appellera modèle minimal.

Proposition 5 *La fonction H qui à toute formule normale doxastique positive (w.r.t.*

[7] $a\bar{b}$ désigne l'état où a est vrai et b est faux.

équivalence), associe son modèle (OCF) minimal, est bijective .

Cette propriété est primordiale pour la suite pour les raisons suivantes :

1. Elle établit une caractérisation sémantique des formules doxastiques positives qui traduisent tout ce que l'agent croit (à tous les degrés) à un moment donné. Elle établit aussi une caractérisation syntaxique (par une formule positive) de tout modèle.
2. Cette correspondance permettra de passer du cadre syntaxique au cadre sémantique (ou vice-versa) en appliquant la fonction H (ou sa réciproque $G = H^{-1}$). Les deux cadres ont néanmoins leur utilité. Le cadre sémantique est plus puissant pour définir les règles de révisions et de mises à jour des croyances (c'est d'ailleurs celui qui est le plus utilisé dans la littérature). Le cadre syntaxique, lui, permet une représentation plus compacte des croyances, il est, computationnellement parlant, plus efficace que le cadre sémantique.

Nous noterons désormais $\mathcal{B}_S$ l'ensemble des états de croyances. Un état de croyances peut être représenté soit par une OCF κ, soit par une formule NDP Φ.
Nous pouvons remarquer, cependant, que les formules NDPs ($\Phi = \mathbf{B}^{\infty}\varphi_{\infty} \wedge \mathbf{B}^{\mathbf{n}}\varphi_n \wedge \ldots \wedge \mathbf{B}^{\mathbf{1}}\varphi_1$) n'ont pas un modèle *unique*. En effet, tout κ supérieur ou égal à $H(\Phi)$ est aussi un modèle de Φ. Ceci vient du fait que les formules NDPs n'expriment que les croyances positives de l'agent, alors qu'un modèle (une OCF) véhicule davantage d'information et notamment tout ce que l'agent ne sait pas et ne croit pas (à tous les degrés de conviction). Pour obtenir, syntaxiquement, toutes les informations "contenues" dans une OCF, on peut enrichir le langage d'une nouvelle famille de modalités, servant à décrire *tout ce que l'agent croit.*

5 Modalité "je crois seulement"

Levesque, dans [12], a introduit la notion de "only knowing" avec un agent. Cette notion a été revue et étendue à plusieurs agents par Halpern et Lakemeyer dans [6]. Nous l'étendons ici aux croyances graduelles (pour un agent).

5.1 "only knowing"

Le langage $\mathcal{L}(\mathtt{PS}, \{\mathbf{B}\})$ est enrichi d'une modalité $\mathbf{N}$. De plus, la formule $\mathbf{O}\Phi$ est le raccourci de $\mathbf{B}\Phi \wedge \mathbf{N}\neg\Phi$. Où $\mathbf{B}\Phi$ signifie (comme précédemment) "l'agent croit au moins Φ", $\mathbf{N}\Phi$ signifie "l'agent croit au plus $\neg\Phi$" et $\mathbf{O}\Phi$ signifie "l'agent croit *seulement* Φ".
Levesque en donne une sémantique et une axiomatique qu'il prouve adéquate et complète.

5.2 Modalité "croire seulement" pour $\mathsf{KD45_G}$

On enrichit le langage d'une famille $\mathtt{N} = \{\mathbf{N}^{\infty}, \mathbf{N}^{\mathbf{1}}, \mathbf{N}^{\mathbf{2}}, \ldots\}$ et d'une famille $\mathtt{O} = \{\mathbf{O}^{\infty}, \mathbf{O}^{\mathbf{1}}, \mathbf{O}^{\mathbf{2}}, \ldots\}$ dans le même esprit que chez [12, 6]. Où $\mathbf{N}^{\mathbf{i}}\neg\varphi$ signifie : "au degré i, l'agent croit au plus φ", et l'opérateur $\mathbf{O}^{\mathbf{i}}$ est le raccourci de $\mathbf{B}^{\mathbf{i}}\varphi \wedge \mathbf{N}^{\mathbf{i}}\neg\varphi$. L'étude complète de la logique définie à partir du langage $\mathcal{L}(\mathtt{PS}, \mathtt{B} \cup \mathtt{N} \cup \mathtt{O})$ s'éloignant trop de notre propos nous nous contenterons ici de donner quelques propriétés de l'opérateur $\mathbf{O}^{\mathbf{i}}$. Le langage permettant donc d'exprimer ce que l'agent croit et seulement ce qu'il croit est $\mathcal{L}(\mathtt{PS}, \mathtt{B} \cup \mathtt{N} \cup \mathtt{O}) = \mathcal{L}^g_{\mathbf{O}}$.

La vérité d'une formule (doxastiquement interprétable) dans une OCF κ est définie par : $\kappa \models \mathbf{N}^{\mathbf{i}}\neg\varphi \Leftrightarrow \forall s(s \models \varphi \Rightarrow \kappa(s) < i)$ et, puisque $\kappa \models \mathbf{B}^{\mathbf{i}}\varphi \Leftrightarrow \forall s(\kappa(s) < i \Rightarrow s \models \Phi)$ on a, pour $\mathbf{O}^{\mathbf{i}}\kappa \models \mathbf{O}^{\mathbf{i}}\varphi \Leftrightarrow \forall s(s \models \Phi \Leftrightarrow \kappa(s) < i)$.

$\kappa \models \mathbf{O^i}\varphi$ est vraie dès que tous les états exceptionnels au moins au degré i sont *les seuls* à satisfaire $\neg\varphi$, ou bien, de manière équivalente, les états exceptionnels au plus au degré $i-1$ sont *les seuls* à satisfaire φ. En particulier $\mathbf{O^1}\varphi$ est satisfaite quand tous les états "normaux" satisfont φ, et que ce sont les seuls.
Notons que $\mathbf{O^1}\top$ signifie que l'agent ne croit rien au degré 1, par conséquent, rien non plus au degré 2, etc. Le seul OCF κ satisfaisant $\mathbf{O^1}\top$ est κ_{vide}.

Proposition 6 *Pour tout φ, les formules suivantes sont des théorèmes :*

1. $\mathbf{O^i}\varphi \rightarrow \mathbf{B^i}\varphi$
2. $\mathbf{O^i}\varphi \leftrightarrow \mathbf{B^i}\varphi \wedge \bigwedge_{s\models\varphi} \neg\mathbf{B^i}(\varphi \wedge \neg Form(s))$

Exemple 3 *Soit* $\mathtt{PS} = \{a, b\}$ *et κ définie par*

$$\kappa = \begin{bmatrix} ab & : & 1 \\ a\bar{b} & : & 0 \\ \bar{a}b & : & 1 \\ \bar{a}\bar{b} & : & \infty \end{bmatrix}.$$

Alors

- $\kappa \models \mathbf{O^1}(a \wedge \neg b)$*, car* $a \wedge \neg b$ *est satisfaite seulement par tous les états normaux.*
- $\kappa \models \mathbf{O^\infty}(a \vee b)$.
- $\kappa \models \mathbf{B^1}a$*, mais* $\kappa \not\models \mathbf{O^1}a$

Définition 7 *Pour toute formule normale doxastique positive (NDP)* $\Phi = \mathbf{B^\infty}\varphi_\infty \wedge \mathbf{B^n}\varphi_n \wedge \ldots \wedge \mathbf{B^1}\varphi_1$, $\mathbf{Only}(\Phi)$ *est la formule*

$$\mathbf{O^\infty}\varphi_\infty \wedge \mathbf{O^{n+1}}\varphi_\infty \wedge \mathbf{O^n}\varphi_n \wedge \ldots \wedge \mathbf{O^1}\varphi_1.$$

Ces formules expriment la totalité des croyances d'un agent ; elles sont satisfaites dans l' *unique* OCF $\kappa_\Phi = H(\Phi)$ définie selon la définition 5. La propriété suivante nous assure de l'existence d'une correspondance bi-univoque entre les NDPs et les OCFs (w.r.t. équivalence dans $\mathsf{KD45_G}$) ainsi qu'une équivalence d'expressivité.

Proposition 7 *Pour toute formule NDP* $\Phi = \mathbf{B^\infty}\varphi_\infty \wedge \mathbf{B^n}\varphi_n \wedge \ldots \wedge \mathbf{B^1}\varphi_1$,

$$\kappa \models \mathbf{Only}(\Phi) \Leftrightarrow \kappa = \kappa_\Phi$$

6 Travaux connexes

La logique possibiliste. La sémantique fournie par les OCFs est très proche de celle des distributions de possibilités (voir [1] pour plus de détails à ce sujet). Du point de vue syntaxique notre langage est plus complet que celui de la logique possibiliste puisqu'il permet, par exemple, de raisonner sur les croyances de croyances.
Par ailleurs, un langage de logique modale avec une sémantique basée sur les distributions de possibilités est proposé dans [3]. Il se rapproche du notre mais lui nonplus ne considère pas les croyances sur les croyances.

Des croyances graduelles basées sur un comptage des mondes. Dans [16], Meyer et van der Hoek développent des modalités graduelles en logique épistémique, le système $Gr(S5)$.
Dans leur sémantique, la croyance en un fait "φ" est d'autant plus forte que le nombre de mondes du modèle satisfaisant φ est grand. La modalité $\mathbf{B^i}$ exprimant le degré de croyance ne possède alors pas les mêmes propriétés que la nôtre et s'applique donc à des exemples bien différents des nôtres.

L'approche de Hans van Ditmarsch. Dans "Prolegomena to dynamic logic for belief revision" [17] van Ditmarsch développe un modèle (de Kripke) $M = \langle S, <, V \rangle$ basé sur la fonction « $<$ » qui à chaque s de S associe une relation $<^s$ de $S \times S$ dite relation de "plausibilité" dans l'état s. Les modalités $\mathbf{B^i}$ qu'il utilise sont identiques aux nôtres et ont les mêmes propriétés. Ces résultats furent établis en parallèle des nôtres [10, 9].

Références

[1] D. Dubois and H. Prade. A synthetic view of belief revision with uncertain inputs in the framework of possibility theory. *Int. J. of Approximate Reasoning*, 17(2–3) :295–324, 1997.

[2] R. Fagin, J. Halpen, Y. Moses, and M. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.

[3] L. Fariñas del Cerro and A. Herzig. Modal logics for possibility theory. In *Proceedings of the First International Conference on the Fundamentals of AI Research (FAIR'91)*. Springer Verlag, 1991.

[4] N. Friedman and J.Y. Halpern. A knowledge based frameword for belief change part I : foundations. In R. Fagin, editor, *Proceedings of the Fifth Conference on Theoretical aspects of reasoning about knowledge*, pages 44–64, 1994.

[5] J. Y. Halpern. Should knowledge entail belief ? *Journal of Philosophical Logic*, 25(5) :483–494, 1996.

[6] Joseph Y. Halpern and Gerhard Lakemeyer. Multi-agent only knowing. In Yoav Shoham, editor, *Theoretical Aspects of Rationality and Knowledge : Proceedings of the Sixth Conference (TARK 1996)*, pages 251–265. Morgan Kaufmann, San Francisco, 1996.

[7] S. Kraus and D. Lehmann. Knowledge, belief, and time. *Theoretical Computer Science*, 58 :155–174, 1988.

[8] P. Lamarre and Y. Shoham. Knowledge, certainty, belief and conditionalisation. In *Proceedings of the* 4^{th} *International Conference on Principles of Knowledge Representation and Reasoning (KR'94)*, pages 415–424, 1994.

[9] N. Laverny and J. Lang. From knowledge-based programs to graded belief based programs, part I : on-line reasoning. *Synthese*.

[10] N. Laverny and J. Lang. From knowledge-based programs to graded belief based programs, part I : on-line reasoning. In *Proceedings of ECAI-04*, pages 368–372, 2004.

[11] W. Lenzen. Recent work in epistemic logic. *Acta Philosophica Fennica*, 30 :1–129, 1978.

[12] Hector J. Levesque. All i know : a study in autoepistemic logic. *Artificial Intelligence*, 42(2-3) :263 – 309, 1990.

[13] J.-J. Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science*. Number 41 in Cambridge Tracks in Theoretical Computer Science, 1995.

[14] Y. Moses and Y. Shoham. Belief as defeasible knowledge. *Artificial Intelligence*, 64(2) :299–322, 1993.

[15] W. Spohn. Ordinal conditional functions : a dynamic theory of epistemic states. In William L. Harper and Brian Skyrms, editors, *Causation in Decision, Belief Change and Statistics*, volume 2, pages 105–134. Kluwer Academic Pub., 1988.

[16] W. van der Hoek and J.-J.Ch. Meyer. Graded modalities for epistemic logic. *Logique et Analyse*, 133-134 :251–270, 1991.

[17] H. van Ditmarsch. Prolegomena to dynamic belief revision. Technical report, University of Otago, New Zealand, 2004.

[18] G.H. von Wright. *An essay in modal logic*. Nortf-Holland, 1953.

[19] F. Voorbraak. Generalized kripke models for epistemic logic. In Y. O. Moses, editor, *Theoretical Aspects of Reasoning about knowledge : Proc. 4th Conference*, pages 214–218. Morgan Kaufman, 1992.